

RESULT 2  
Q4P7C4 USTMA  
ID Q4P7C4 USTMA PRELIMINARY;  
PRT; 620 AA.







RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahay J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S., Krzywiński M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RL [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Embryo;  
RC Klein S., Gerhard D.S.;  
RL Submitted (OCT-2004) to the EMBL/GenBank/DBSJ databases.  
RL EMBL: BC084467; AAH84467.1; -; mRNA.  
DR GO: GO:0005764; C:lysosome; IEA.  
DR GO: GO:0016020; C:membrane; IEA.  
DR GO: GO:0004872; F:receptor activity; IEA.  
DR GO: GO:0005215; F:transporter activity; IEA.  
DR GO: GO:0006810; P:transport; IEA.  
DR InterPro: IPR000479; C:IMR.  
DR Pfam: PF00878; C:IMR; 1.  
FT NON\_TER 289  
FT SEQUENCE 289 AA; 32022 MW; 05950766D19DC43A CRC64;  
  
Query Match 52.2%; Score 109; DB 2; Length 289;  
Best Local Similarity 59.5%; Pred. No. 0.098;  
Matches 25; Conservative 1; Mismatches 16; Indels 0; Gaps  
  
Qy 4 AAAAAAEEEEAAAAAEEEEAAAAKKKKKKKKKKKKKKKKKKKKKK 45  
||| : |  
Db 233 AAALNSTANERLVIRVYETDTDTKKKKKKKKKKKKKKKKKKKK 274  
  
RESULT 8  
Q871E9\_NEUCR  
ID Q871E9\_NEUCR PRELIMINARY; PRT; 1240 AA.  
AC Q871E9;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hypothetical protein B7H23.320.  
DE Names=B7H23.320;  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,  
RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.  
RL [2]  
RP NUCLEOTIDE SEQUENCE.  
RL German Neurospora genome project;  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.  
DR ENBL: BX294026; CAD71057.1; -; Genomic\_DNA.  
DR InterPro: IPR001611; LRR.  
DR Pfam: PF00560; LRR 1; 1.  
DR PRINTS; PR00019; LEURICHRP.  
KW Hypothetical protein; Leucine-rich repeat; Repeat.  
SQ SEQUENCE 1240 AA; 134655 MW; 37660713FBA9CD58 CRC64;  
  
Query Match 52.2%; Score 109; DB 2; Length 1240;  
Best Local Similarity 57.9%; Pred. No. 0.35;  
Matches 22; Conservative 10; Mismatches 6; Indels 0; Gaps

```

Qy 8 EAAAAAEEAAAAAEEAAAAKKKKKKKKKKKKKKKKKKKKKKKK 45
Db 307 ESSREAEAAAAAEEAAAAETKKKKKKKKSEERKARKEKK 344

RESULT 9
Q6AWH4 DROME
ID Q6AWH4 DROME PRELIMINARY; PRT; 537 AA.
AC Q6AWH4;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE LD17709p (Fragment).
GN Names=CG3654;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Phyllophoridae; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RA Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Celniker S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL; BT015274; AAT94503.1; -, mRNA.
DR InterPro; IPR000637; A-T hook.
DR PRINTS; PR00929; ATHOOK.
FT NON TER 537
SQ SEQUENCE 537 AA; 57754 MW; C4751515B16ABA3E CRC64;

Query Match 51.9%; Score 108.5; DB 2; Length 537;
Best Local Similarity 50.9%; Pred.No.0.18;
Matches 28; Conservative 3; Mismatches 13; Indels 11; Gaps
9

Qy 2 EAAAAAEEAAAAAEEAAAAAEEAAAAKK-----KKKKKKKKKKKKKKKKKK 45
Db 483 KCASPKKGGAPAPAPARNKAAAKLDRKEIKDHKKKKKKKKKKKKKK 537

RESULT 10
I7DS295 XENLA
ID I7DS295 XENLA PRELIMINARY; PRT; 143 AA.
AC I7DS295;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole.
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

```





```
RESULT 15
Q7SD04 NEUCR
ID Q7SD04 NEUCR PRELIMINARY; PRT; 94 AA.
AC Q7SD04
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted protein.
DE Name=NCU08122.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysseelis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000069; EAA34645.1; -; Genomic_DNA.
DR GO; GO:0005694; C:chromosome; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006334; P:nucleosome assembly; IEA.
DR InterPro; IPR005819; Histone H5.
DR PRINTS; PR00624; HISTONEH5.
SQ SEQUENCE 94 AA; 11689 MW; 3039FBD548B2C683 CRC64;

Query Match 50.2%; Score 105; DB 2; Length 94;
Best Local Similarity 59.0%; Pred. No. 0.079;
Matches 23; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 AAAAAAAAAAAAAAAAAAKKKKKKKKKKKKKKKKKKKKK 45
Dd 2 AEKVVVVALCGRKEEKQGEKKKKKKKKKKKKKKKKKKKK 40

Search completed: January 4, 2006, 10:22:04
Job time : 237 secs
```



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	103	49.3	248	1	HSUR1P	histone H1, gonada	
2	102	48.8	517	2	T49173	hypothetical prote	
3	101.5	48.6	442	2	T39683	zootin-like protei	
4	100	47.8	215	2	T32683	nucleoporin p62 ho	
5	100	47.8	380	2	T45395	hypothetical prote	
6	98	46.9	206	2	S09388	histone H1 - sea u	
7	95	45.9	433	2	S25194	zootin - yeast (Sa	
8	95	45.5	166	2	T18513	hypothetical prote	
9	91	43.5	229	2	JC7219	nuclear protein SR	
10	90.5	43.3	383	2	AC8315	F2H15.19 protein -	
11	90	43.1	153	2	S59591	histone H2B (clone	
12	90	43.1	483	2	F71619	hypothetical prote	
13	90	43.1	529	2	T50609	hypothetical prote	
14	89	42.6	208	2	T23778	histone H1.1 - Cae	
15	89	42.6	285	2	S19113	cgr-4 protein - C	
16	88	42.1	392	2	T17555	hypothetical prote	
17	88	42.1	409	2	T18726	hypothetical prote	
18	87	41.6	107	2	C86477	protein F1504.29 [	
19	87	41.6	565	2	T47775	hypothetical prote	
20	86	41.1	347	2	E85525	Tola protein PA097	
21	86	41.1	455	2	A89193	protein B205.10 [	
22	85.5	40.9	1280	2	G96796	hypothetical prote	
23	85	40.7	102	2	D64363	ribosomal protein	
24	85	40.7	153	2	S59587	histone H2B (clone	
25	85	40.7	241	2	JN0748	histone H1-II - Vo	
26	85	40.7	560	2	T06377	SAR DNA-binding pr	
27	85	40.7	849	2	T01286	probable RNA-bindi	
28	84.5	40.4	228	2	T36379	probable peptidase	
29	84	40.2	409	2	T24543	hypothetical prote	

```
C:Genetics:
A:Gene: ATSP:T20N10.250
A:Map position: 3
A:Introns: 312/3; 359/3; 444/3
C:Superfamily: Arabidopsis thaliana hypothetical protein F17J16.30

Query Match      48.8%; Score 102; DB 2; Length 517;
Best Local Similarity 80.8%; Pred. No. 0.18;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 20 EAAAAAKKKKKKKKKKKKKKKKKKKKKKKK 45
Db 440 ERVGKKKKKKKKKKKKKKKKKKKKKKKK 465

RESULT 3
T39683
zootin-like protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C:Accession: T39683; T40195
R:Oliver, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z21869
A:Accession: T39683
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-124 <OLI>
A:Cross-references: UNIPROT:Q9Y718; UNIPARC:UPI00000697C4; EMBL:AL049489; PIDN:CAB39796.
A:Experimental source: strain 972h-; cosmid c1778
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, February 1998
A:Reference number: Z21910
A:Accession: T40195
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 89-442 <WOO>
A:Cross-references: UNIPARC:UPI000006AD4A; EMBL:Z97992; PIDN:CAB10796.1; GSPDB:GN00067;
A:Experimental source: strain 972h-; cosmid c30D10
C:Genetics:
A:Gene: SPDB:SPBC1778.01c; SPDB:SPBC30D10.01
A:Map position: 2

Query Match      48.6%; Score 101.5; DB 2; Length 442;
Best Local Similarity 56.0%; Pred. No. 0.18;
Matches 28; Conservative 3; Mismatches 12; Indels 7; Gaps 1;

QY 2 EAAAAA-----EAAAAAFAAAAAAFAAAAAKKKKKKKKKKKKKKKKKKKKKKKK 44
Db 309 EAAAAAQQKKEERRAAEEAAKASAAAAANKKAKEDKKKQKRDKKVVK 358

RESULT 4
I52523
nucleoporin p62 homolog - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 05-Nov-1999
C:Accession: I52523
R:Wang, Z.Q.; Akmal, K.M.; Kim, K.H.
Biol. Reprod. 51, 1022-1030, 1994
A:Title: An unusual nucleoporin-related messenger ribonucleic acid is present in the ger
A:Accession: I52523
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: mRNA
A:Residues: 1-215 <RES>
A:Cross-references: UNIPARC:UPI000005E5FE8; GB:S75997; NID:G913245; PIDN:AAB33384.1; PID:
A:Experimental source: testis

Query Match      47.8%; Score 100; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 26 KKKKKKKKKKKKKKKKKKKKKKKKK 45
Db 35 KKKKKKKKKKKKKKKKKKKKKKKKK 54

RESULT 5
T46395
hypothetical protein DKFp434I1120.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46395
R:Ottewaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23031
A:Accession: T46395
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-380 <AAA>
A:Cross-references: UNIPROT:Q9NT34; UNIPARC:UPI000006D74E; EMBL:AL137556
A:Experimental source: adult testis; clone DKFp434I1120
C:Genetics:
A:Note: DKFp434I1120.1

Query Match      47.8%; Score 100; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 KKKKKKKKKKKKKKKKKKKKKKKKK 45
Db 355 KKKKKKKKKKKKKKKKKKKKKKKKK 374

RESULT 6
S09388
histone H1 - sea urchin (Parechinus angulosus)
C:Species: Parechinus angulosus (angulate urchin)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S09388
R:Hill, C.S.; Martin, S.R.; Thomas, J.O.
EMBO J. 8, 2591-2599, 1989
A:Title: A stable alpha-helical element in the carboxy-terminal domain of free and chrom
A:Reference number: S09388; MUID:90060019; PMID:2583125
A:Accession: S09388
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-206 <HIL>
A:Cross-references: UNIPROT:Q7M409; UNIPARC:UPI0000177176
C:Superfamily: histone H1
C:Keywords: chromosomal protein

Query Match      46.9%; Score 98; DB 2; Length 206;
Best Local Similarity 55.8%; Pred. No. 0.19;
Matches 24; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY 2 EAAAAAFAAAAAFAAAAAAFAAAAAKKKKKKKKKKKKKKKKKKKKKKKK 44
Db 140 KAAAKRKAALAKKAAAKRKAATKAKKAKPKKTKTAAKAKK 182

RESULT 7
S25194
zootin - yeast (Saccharomyces cerevisiae)
N:Alternate names: probable Z-DNA-binding protein; protein G9554; protein YGR285c
C:Species: Saccharomyces cerevisiae
C:Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 09-Jul-2004
C:Accession: S25194; S64620; S19066
R:Zhang, S.; Lockshin, C.; Herbert, A.; Winter, E.; Rich, A.
EMBO J. 11, 3787-3796, 1992
A:Title: Zootin, a putative Z-DNA binding protein in Saccharomyces cerevisiae.
A:Reference number: S25194; MUID:93010971; PMID:1396572
A:Accession: S25194
A:Molecule type: DNA
A:Residues: 1-433 <ZHA>
```

hypothetical protein PFB0235w - malaria parasite (Plasmodium falciparum)

```
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C:Accession: F71619
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
.; Perera, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: F71619
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-483 <GAR>
A:Cross-references: UNIPROT:O96148; UNIPARC:UPI000017B604; GB:AE001362; NID
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0235W

Query Match 43.1%; Score 90; DB 2; Length 483;
Best Local Similarity 85.0%; Pred. No. 1.5;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 26 KKKKKKKKKKKKKKKKKKK 45
Db 449 KKKKKKKKKKKKKKKKKKK 468

RESULT 13
T50609
Hypothetical protein DKFp761B2423.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: T50609
R:Bloescker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, June 2000
A:Reference number: Z25143
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-529 <AAA>
A:Cross-references: UNIPROT:Q9NPX4; UNIPARC:UPI000006DB14; EMBL:AL359564
A:Experimental source: adult amygdala; clone DKFp761B2423
C:Genetics:
A:Note: DKFp761B2423.1

Query Match 43.1%; Score 90; DB 2; Length 529;
Best Local Similarity 45.0%; Pred. No. 1.6;
Matches 18; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 5 AAAAAAAAEAAAAAEAAAAKKKKKKKKKKKKKKKKKK 44
Db 444 SGGQTRSHSSASSAESQDSKKKKKKKKKKKKKKKK 483

RESULT 14
T23778
histone H1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23778; T42231; S09130; S01817
R:Percy, C.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19798
A:Accession: T23778
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-208 <WLD>
A:Cross-references: UNIPROT:Q93901; UNIPARC:UPI00000778B5; EMBL:Z79603; PIDN:CAB01892.1;
A:Experimental source: clone M163
R:Jedrussik, M.; Schulze, E.
submitted to the EMBL Data Library, August 1997
A:Description: The histone H1 complement of Caenorhabditis elegans.
A:Reference number: Z22091
A:Accession: T42231
```

```
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-208 <JED>
A:Cross-references: UNIPARC:UPI00000778E5; EMBL:AF017810; PIDN:AAB70665.1
R:Sanicola, M.; Ward, S.; Childs, G.; Emmons, S.W.
J. Mol. Biol. 212, 259-268, 1990
A:Title: Identification of a Caenorhabditis elegans histone H1 gene family. Characterizat
A:Reference number: S09130; MUID:90204554; PMID:1969492
A:Accession: S09130
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-43,'T','45-83,'H','85-100,'R','102-208 <SAN>
A:Cross-references: UNIPARC:UPI0000168900; GB:X53277; NID:g10885; PIDN:CAA37372.1; PID:g
R:Vanlenteren, J.R.; van Bun, S.M.; van Beumen, J.J.
Biochem. J. 255, 647-652, 1988
A:Title: The primary structure of the major isoform (H1.1) of histone H1 from the nematoc
A:Reference number: S01817; MUID:89076229; PMID:3202838
A:Accession: S01817
A:Molecule type: protein
A:Residues: 2-43,'T','45-100,'K','102-208 <VAN>
A:Cross-references: UNIPARC:UPI00001718E
C:Genetics:
A:Gene: CESP:M163.3; his-24
A:Map position: X
A:Introns: 79/3
C:Superfamily: histone H1
C:Keywords: blocked amino end; chromosomal protein; DNA binding; nucleosome; nucleus
F:2-208/Product: histone H1.1 #status predicted <MAT>
F:2/Modified site: blocked amino end (Ser) (in mature form) (probably acetylated) #status

Query Match 42.6%; Score 89; DB 2; Length 208;
Best Local Similarity 53.5%; Pred. No. 0.96;
Matches 23; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

QY 2 EAAAAAEAAAAAEAAAAKKKKKKKKKKKKKKKKKK 44
Db 113 EKAATAKKPAAAKPAAAKPAAAKKATGEKKAKKPAAPKPK 155

RESULT 15
S19113
cgcr-4 protein - Chlamydomonas reinhardtii (fragment)
C:Species: Chlamydomonas reinhardtii
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S19113; S14466
R:Wakarchuk, W.W.; Mueller, F.W.; Beck, C.F.
Plant Mol. Biol. 18, 143-146, 1992
A:Title: Two GC-rich DNA elements of Chlamydomonas reinhardtii with complex arrangements
A:Reference number: S19113; MUID:92119224; PMID:1731966
A:Accession: S19113
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <WAK>
A:Cross-references: UNIPROT:Q39598; UNIPARC:UPI00000A66FD; EMBL:X17208; NID:g18136; PIDN:
C:Genetics:
A:Gene: cgcr-4

Query Match 42.6%; Score 89; DB 2; Length 265;
Best Local Similarity 68.8%; Pred. No. 1.1;
Matches 22; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 AAAAAAEAAAAAEAAAAKKKKKKKKKKKK 34
Db 176 AAEEAFAAAAAAEAAAAKARAAAEAKAEADK 207

Search completed: January 4, 2006, 10:22:47
Job time : 40 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 09:57:22 ; Search time 185 Seconds  
(without alignments)  
106.876 Million cell updates/sec

Title: US-09-461-684C-4  
Perfect score: 209  
Sequence: 1 CEAATAAEAAAAAEEAAAAE.....KKKKKKKKKKKKKKKKKKK 45

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001s.\*
- 5: Geneseq2002s.\*
- 6: Geneseq2003s.\*
- 7: Geneseq2003bs.\*
- 8: Geneseq2004s.\*
- 9: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	209	100.0	45	3 AAB13783	Aab13783 Soluble t
2	119	56.9	36	3 AAB23585	Aab23585 Agk21 lin
3	119	56.9	211	8 ADH89966	Adh89966 Synthetic
4	119	56.9	630	3 AAB23591	Aab23591 Modified
5	119	56.9	640	3 AAB23593	Aab23593 Modified
6	118	56.5	2451	4 ABB71574	Abb71574 Drosophila
7	115.5	55.3	123	4 AAO01368	Aao01368 Human pol
8	115	55.0	113	4 AAO03703	Aao03703 Human pol
9	113	54.1	122	5 ABP66694	Abp66694 Human bre
10	111	53.1	63	4 AAU18275	Aau18275 Novel hum
11	111	53.1	63	5 ABG92696	Abg92696 Human DNA
12	111	53.1	63	7 ADC25413	Adc25413 Human ext
13	110	52.6	141	4 ABG26718	Abg26718 Novel hum
14	109	52.2	25	3 AAB13781	Aab13781 Soluble p
15	109	52.2	59	3 AAY98495	Aay98495 Nuclear l
16	109	52.2	59	3 AAY59040	Aay59040 Nuclear l
17	109	52.2	59	4 AAB45848	Aab45848 Nucleic a
18	109	52.2	59	4 AAU04285	Aau04285 Nuclear l
19	109	52.2	74	4 AAO03278	Aao03278 Human pol
20	109	52.2	75	4 AAO08943	Aao08943 Human pol
21	109	52.2	112	4 AAO02477	Aao02477 Human pol
22	109	52.2	150	4 AAU18240	Aau18240 Novel hum
23	109	52.2	150	5 ABG92661	Abg92661 Human DNA
24	109	52.2	150	7 ADC25378	Adc25378 Human ext

ALIGNMENTS

RESULT 1  
AAB13783  
ID AAB13783 standard; peptide; 45 AA.  
AC AAB13783;  
XX  
XX  
DT 10-NOV-2000 (first entry)  
DE Soluble tandem pEA/ pK peptide conjugate.  
DE  
DE  
KW pk peptide; cytostatic; vaccine; cytotoxic T cell; CTL; immunotherapy;  
KW major histocompatibility complex class I; MHC class I; antigen; tumour;  
KW prostate; breast; multiple myeloma; pEA peptide.  
XX  
XX  
OS Unidentified.  
XX  
PN WO200035949-A1.  
XX  
PD 22-JUN-2000.  
XX  
PF 14-DEC-1999; 99WO-US029724.  
XX  
PR 14-DEC-1998; 98US-0112324P.  
XX  
XX (DEND-) DENDREON CORP.  
XX  
XX Laus R, Hakim I, Vidovic D;  
XX  
XX WPI; 2000-442365/38.  
XX  
XX Antigen modified by the covalent addition of a peptide that facilitates entry into antigen presenting cells, useful for producing compositions for immunizing against tumors and pathogens.  
XX  
XX Claim 2; Page 26; 34pp; English.  
XX  
XX The present invention relates to compositions of modified soluble protein antigens capable of eliciting an enhanced in vivo cytotoxic T cell (CTL) response i.e. a major histocompatibility complex (MHC) class I molecule peptide sequence which facilitate entry of the antigen into antigen presenting cells (APCs). The present sequence is one such peptide sequence which can be used to modify the soluble antigens. The present sequence is tandem pEA/ pK peptide conjugate. The modified antigen composition may be used for immunising against, or treating a tumour e.g. prostate and breast carcinoma or multiple myeloma, or pathogen in mammals

25 108 51.7 50 5 AAU69736 Aau69736 Cell deat  
26 108 51.7 126 4 AAO12105 Aao12105 Human pol  
27 108 51.7 126 4 AAO02946 Aao02946 Human pol  
28 107 51.2 48 4 AAO08707 Aao08707 Human pol  
29 107 51.2 63 4 AAU18192 Aau18192 Novel hum  
30 107 51.2 63 5 ABG92613 Abg92613 Human DNA  
31 107 51.2 63 7 ADC25330 Adc25330 Human ext  
32 107 51.2 83 4 AAO02961 Aao02961 Human pol  
33 106 50.7 26 4 AAO08995 Aao08995 Human pol  
34 106 50.7 57 4 ABG26720 Abg26720 Novel hum  
35 106 50.7 64 4 AAO03024 Aao03024 Human pol  
36 106 50.7 74 4 AAU18205 Aau18205 Novel hum  
37 106 50.7 74 4 AAU18277 Aau18277 Novel hum  
38 106 50.7 74 5 ABG92698 Abg92698 Human DNA  
39 106 50.7 74 5 ABG92636 Abg92636 Human DNA  
40 106 50.7 74 7 ADC25343 Adc25343 Human ext  
41 106 50.7 74 7 ADC25415 Adc25415 Human ext  
42 106 50.7 128 3 AAY86248 Aay86248 Human sec  
43 106 50.7 128 6 ABO53391 Abo53391 Novel hum  
44 106 50.7 272 4 AAU23799 Aau23799 Novel hum  
45 105 50.2 27 3 AAB59105 Aab59105 Breast an









DE Human breast specific protein SEQ ID NO 196.  
XX Human; breast; cytostatic; cancer; transgenic; gene therapy; vaccine.  
XX Homo sapiens.  
XX WO200266605-A2.  
XX 29-AUG-2002.  
XX 14-FEB-2002; 2002WO-US004284.  
XX 15-FEB-2001; 2001US-0268999P.  
XX (DIAD-) DIADEXUS INC.  
XX Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferkey R;  
PI Sun Y, Liu C;  
XX WPI; 2002-713345/77.  
XX New isolated breast specific nucleic acid molecules and polypeptides,  
PT useful for identifying, diagnosing, monitoring, staging, imaging and  
PT treating breast cancer and non-cancerous disease states in breast tissue.  
XX Claim 11; Page 246; 254pp; English.  
XX The invention relates to human breast specific nucleic acids (I)  
CC comprising: (a) a sequence encoding any one of 95 protein sequences  
CC (ABP66614-ABP66708); (b) any one of 115 polynucleotide sequences  
CC (ABV83558-ABV83672); (c) a molecule that selectively hybridizes to (a) or  
CC (b); (d) a molecule having at least 60% sequence identity to (a) or (b).  
CC The breast specific nucleic acid molecules, polypeptides and antibodies  
CC are useful for identifying, diagnosing, monitoring, staging, imaging and  
CC treating breast cancer and non-cancerous disease states in breast tissue.  
CC They are also useful for producing transgenic animals and cells and  
CC producing engineered breast tissue for treatment and research. The  
CC transgenic animals are useful as animal model systems used in elaborating  
CC the biological function of the polypeptides, studying conditions and/or  
CC disorders associated with aberrant expression and in screening for  
CC compounds effective in ameliorating the conditions. The polynucleotides  
CC are useful for gene therapy and in vaccines  
XX SQ Sequence 122 AA;  
Query Match 54.1%; Score 113; DB 5; Length 122;  
Best Local Similarity 88.5%; Pred. No. 0.00043;  
Matches 23; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 20 EAAAAAKKKKKKKKKKKKKKKKKKKKK 45  
Db 2 DAARAGKKKKKKKKKKKKKKKKKKKK 27  
RESULT 10  
AAU18275  
ID AAU18275 standard; protein; 63 AA.  
XX AAU18275;  
AC AAU18275;  
XX 21-NOV-2001 (first entry)  
DT Novel human DNA-binding protein #122.  
DE Human; DNA-binding protein; histone; chromo domain protein;  
XX chromatin organisation modifier; Y-box binding protein; DNA organisation;  
KW gene transcription; malignant disease; autoimmune disorder;  
KW rheumatic disease; genetic abnormality; infectious disease;  
KW neurological disorder; gene therapy; immunomodulatory; anti-HIV;  
KW anti rheumatic; anti microbial; cytostatic.  
XX Homo sapiens.  
OS XX

PN WO200155162-A1.  
XX 02-AUG-2001.  
XX 17-JAN-2001; 2001WO-US001305.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225577P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.





PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0233080P.  
PR 08-SEP-2000; 2000US-0233081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2000US-0254097P.  
PR 17-JAN-2001; 2001US-0259678P.  
XX 17-JAN-2001; 2001US-00764846.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Ruben SM, Barash SC;  
XX  
XX WPI; 2003-605749/57.  
DR N-PSDB; ABC25285.  
XX  
PT New DNA-binding proteins and gene encoding them, useful for diagnosing,  
PT treating and/or preventing e.g. neurological, inflammatory, infectious,  
PT cardiovascular, autoimmune, respiratory, neoplastic or digestive  
XX diseases.  
XX  
PS Claim 11; SEQ ID NO 260; 226pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (cDNA)  
CC encoding a human extracellular matrix protein, representing one of 161  
CC novel genes. Also included are recombinant vectors, host cells  
CC (expressing the protein), the extracellular matrix proteins (including  
CC their fragments, epitopes and homologues), an isolated antibody that  
CC binds specifically to the protein, diagnosing a pathological condition or  
CC susceptibility to a pathological condition (comprising determining the  
CC presence or absence of a mutation in the nucleic acid and diagnosing a  
CC condition based on the presence or absence of the mutation), diagnosing a  
CC pathological condition or susceptibility to a pathological condition  
CC (comprising determining the presence or amount of expression of the  
CC protein in a biological sample and diagnosing a condition based on the  
CC presence or amount of expression of the protein), preventing, treating or  
CC ameliorating a medical condition by administering the nucleic acid or  
CC protein to a mammalian subject, identifying a binding partner to the  
CC protein, the gene corresponding to the cDNA sequence, and identifying an



AA98495  
ID AA98495 standard; peptide; 59 AA.  
XX  
AC AA98495;  
XX  
DT 31-JUL-2000 (first entry)  
XX  
DE Nuclear ligand used in nucleic acid transporter system.  
XX  
KW Transporter system; nucleic acid delivery; gene therapy; cancer;  
KW carcinogenesis; cardiovascular disease; infection.  
XX  
OS Synthetic.  
XX  
PN US6033884-A.  
XX  
PD 07-MAR-2000.  
XX  
PF 14-DEC-1993; 93US-00167641.  
XX  
PR 20-MAR-1992; 92US-00855389.  
PR 19-MAR-1993; 93WO-US002725.  
XX  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
XX  
PI Gottchalk S, Sparrow J, Cristiano RJ, Woo SLC, Smith LC;  
XX WPI; 2000-281993/24.  
DR  
XX  
PT System for transporting nucleic acid into cells, useful e.g. in gene  
PT therapy and for generating transgenic animals, comprises binding agent  
PT linked to nucleic acid, surface ligand and lytic agent.  
XX  
PS Claim 16; Col 123-124; 108pp; English.  
XX  
CC The present invention relates to a transporter system for delivering  
CC nucleic acid to a cell. The system comprises a nucleic acid binding  
CC complex, consisting of a binding molecule bonded non-covalently to the  
CC nucleic acid, and covalently to a surface ligand, and a lytic agent. The  
CC binding molecule is spermine or a spermidine derivative. Nucleoside  
CC sequences AA3633-A3652 and peptide sequences AA98456-Y98500 are used  
CC in the construction of the transporter system of the invention. The  
CC transporter system is used in gene therapy, particularly to deliver  
CC nucleic acids to hepatocytes, muscle cells or bone forming cells, e.g. for  
CC treating cardiovascular disease, cancer, and infection. The transporter  
CC systems are also used to create transgenic animals (as models for human  
CC carcinogenesis or disease or for drug testing). Other uses include  
CC transforming cells to produce proteins, or transfecting cells in vitro  
CC to study the function of the nucleic acid. The use of a surface ligand  
CC allows specific targeting of selected cells and tissues. The lytic agent  
CC provides for release of the nucleic acid into the cellular interior, from  
CC endosomes, without requiring endosomal or lysosomal degradation  
XX  
SQ Sequence 59 AA;  
Query Match 52.2%; Score 109; DB 3; Length 59;  
Best Local Similarity 88.5%; Pred. No. 0.00054;  
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 20 EAAAAA  
Db 12 EAPYAKK  
Search completed: January 4, 2006, 10:18:04  
Job time : 187 secs